Exploring the factor structure of measures of confidence in procedural justice and performance of the criminal justice system by actor. A latent-variables approach

Anjuli Van Damme (Ghent University)
Lieven Pauwels (Ghent University)
Nicole Haas (NSCR, Netherlands Institute for the Study of Crime and Law Enforcement)

Abstract

It is commonly acknowledged that many survey measures of confidence in the criminal justice system suffer from a number of measurement problems. A major criticism is the use of single-item questions to evaluate the criminal justice system in general. In this article we contribute to an existing knowledge of measures by assessing two theoretical dimensions of confidence in the criminal justice system: (1) procedural justice or fairness and (2) performance of the criminal justice system. The innovative part of this study is that it explicitly tests to what extent we can measure these dimensions by type of actor (the police, the public prosecutor, judges and the criminal justice system in general). Items that are highly similar in content, but different by type of actor, were presented to a sample of students enrolled at Ghent University. A latent variable modelling approach was used to test the reliability and correlational validity of our multiple-component instrument. Confirmatory factor analyses were conducted on a battery of items that measure procedural justice and performance of three actors within the criminal justice system and the criminal justice system in general. Special attention is paid to the existence of non-random error between indicators that result from question wording. Our analyses suggest that it is possible to distinguish between procedural justice and performance by actor. We tested and found that items that refer to the same actor have correlated error and we propose a way of dealing with it.

Key-words: Reliability, correlational validity, measurement error, attitudes towards the criminal justice system, procedural justice and performance of the criminal justice system, latent variable approach

Introduction

Large-scale surveys of the general population are very popular in the social sciences. In the field of criminology, a growing interest in survey methodology can be observed since the 1950s. One major reason for an increasing body of research in this tradition is the discovery of bias in official measurement instruments; more specifically biases in police statistics. Official statistics tend to underestimate true rates of victimisation in the population, and are said to be seriously biased with respect to race, gender and social class. Another important reason for the wide-spread interest in surveys lies in their potential to serve as a means to empirically test causal theories of offending, victimisation, fear of crime and confidence in the criminal justice system. The research tradition that focuses on confidence in the criminal justice system has experienced a series of problems that are highly similar to those in research on fear of crime. A first lacuna concerns the weak theoretical and conceptual framework surrounding research within this tradition. Obviously, both the policy driven character of the early – and later – large scale confidence surveys, and the positivistic approach of that era, are largely indebted to this. A second lacuna - which in essence has the same roots as the first - is best described as a conservative and more or less outdated methodology and measurement. Because confidence in the criminal justice system has predominantly been measured in many surveys with one single indicator for reasons of 'comparability' (e.g. 'on a scale of 0 to 10, how much confidence do you have in the criminal justice system?'), all claims concerning the reliability and validity can be considered questionable. In recent years, numerous authors have indeed objected to this conservative tenacity, which has resulted in a tendency towards using multi-item measures. These provide a far better way of measuring complex concepts like 'confidence in the criminal justice system'. Currently, when conducting surveys, it is commonly assumed that attitudes such as 'confidence in the criminal justice system' are not directly observable. Such attitudes are latent concepts, which are made observable through the use of indicators. These indicators are assumed to constitute a valid representation of the underlying concept. A careful selection of indicators is therefore essential (Jackson et al. 2011). In designing measures for confidence in the criminal justice system, choices must be made with regard to the wording of questions, the actors that are referred to in the questions, the response scale, the question context and the technique of data collection. Each of these choices and each combination of choices can lead to different errors.

In recent years, some key studies have been published on the measurement of confidence in the criminal justice system. Research consistently shows that confidence is a

multidimensional concept (Hough & Roberts, 2004; Haas, 2010). An important distinction has been made between a sense of justice based on process (confidence in procedural justice of fairness) and one based on the outcome (confidence in the effectiveness of the system). In this contribution we focus on the measurement of these two major dimensions of confidence in the criminal justice system that also play a central role in Tom Tyler's 'procedural justice' theory (Tyler, 2003, 2006, 2007, 2011; Gau et al., 2011). While the measurement instrument is restricted to the measurement of these two aspects of confidence, respondents are asked to answer the items that measure procedural justice and performance using different formulations: as an alternative, we refer in the questionnaire to the police, the public prosecution, judges and lastly, the criminal justice system in general.

Goal of the present study: measuring dimensions of confidence by actor

In criminological inquiries the construction of valid concepts has been called problematic (Pauwels & Pleysier, 2005). Studying the reliability and validity of measurement instruments is necessary to gain insights into the possibilities of a measurement instrument for the purpose of testing theories of confidence. The present study will therefore focus on the reliability and validity of a detailed measurement instrument. The fact that our measurement instrument allows for differentiating between dimensions of confidence namely performance and procedural justice, towards various actors of the criminal justice system yields an interesting question from a methodological point of view: the research question that has to be answered in the present study is: ***"To what extent is it possible to measure procedural justice and performance when changing the reference group (i.e. the police, prosecutors, judges and the criminal justice system in general)".*** To answer this general research question, we will evaluate the factor structure and correlational validity of the concepts of procedural justice of, and performance of, the criminal justice system in relation to these different actors of the criminal justice system and the criminal justice system in general. This is done through a latent variables approach, which allows to model non-random errors that may arise as a consequence of offering similar questions. The advantages of this approach are discussed, after a brief argumentation of why it is necessary –both in criminological inquiries – to pay attention to methodological issues as in substantive research questions. Particular concern arises when detailed questionnaires employ sequences of questions with a common format and highly identical response options to measure attitudes towards different actors of the criminal justice system. Respondents sometimes may interpret a series of questions and

response options in an particular way. In other words, the way a respondent answers the initial item in a survey battery may serve to anchor subsequent answers. This sort of 'response set' implies that the errors of measurement associated with one survey item are not independent of errors associated with others in the same question battery. Statistical procedures that presuppose random measurement error, such as exploratory factor analysis and Likert-scales, may therefore produce misleading results. In that case, the reliability and validity of the measures are at stake. If and only if there are theoretical rationales to suspect non-random errors due to the questionnaire design, the researcher should test such assumptions and model them. Ignoring this structure may offer a misleading account of the structure of attitudes towards the criminal justice system.

Why studying theoretical dimensions of confidence in the criminal justice system?

One of the main problems with current measurements of confidence is a lack of theoretical clarity. Up to now it has remained largely unclear what confidence actually entails. Too often, institutional confidence is a combined scale formed by the summation of the confidence scores on different institutions (e.g. Van Damme et al, 2010). An example is the following bi-annual survey of the Flemish government in which a summation of confidence scores is made on: the Belgian criminal justice system, the Flemish parliament, the Flemish government, the Belgian federal parliament and the Belgian federal government. In the European Social Survey (ESS), the following institutions were included: The Belgian Federal parliament, the legal system, the police, the politicians, the European parliament and the United Nations Organisation (Van de Velde & Pauwels, 2010). These studies indeed show that levels of confidence differ by actor, but this is not helpful if one wants to study confidence in the criminal justice system in great detail. With the aim of improving the measurement of confidence, we need to know exactly what we want to measure. In order to make confidence more tangible and comprehensible, it is useful to start by making a distinction between *the procedural justice model* and *the instrumental model*.

According to the *procedural justice model*, confidence depends largely on perceptions of fairness (Benesh & Howell 2001; Sunshine & Tyler 2003a; Thibaut & Walker 1975; Tyler 2006). *Procedural justice* concerns the integrity and fairness of the justice system (cf. Roberts & Hough 2005; Hough et al. 2010). It constitutes a firm and durable set of attitudes toward the legitimacy of the institution (Caldeira and Gibson, 1995; Reisig et al., 2007). A person

may disagree with the way that an institution deals with a particular issue, yet still accept its overall authority. *Procedural justice or fairness* can for instance be measured by using statements such as "The police are honest and trustworthy" (Sunshine & Tyler, 2003b). Benesh (2006) demonstrated that the more respondents found the court and judges to be fair, the more likely they were to express "a great deal" of confidence in them. Similarly, in a recent study on confidence in the police, which presented respondents with items such as "The police treat people fairly", evidence was found for a strong relationship between perceived fairness of the police and public confidence in it (Jackson & Sunshine, 2007).

According to the *instrumental model*, confidence is developed and maintained through the *performance or effectiveness* of the justice system (Sunshine & Tyler, 2003b). This perceived performance can be seen as an evaluation of how well the system acts, which depends on the extent to which one's expectations of the institution's functioning are met (Caldeira & Gibson, 1995). Confidence in the courts, for instance, is expected to depend on the favourability of the outcome. A civil law study carried out in Scotland shows clear evidence of this effect: 70 percent of successful litigants found the outcome fair, compared to the 10 percent who lost their case (Genn & Paterson, 2001). In the case of the police, confidence could be affected by the extent to which they are able to effectively fight crime and disorder. For instance in a study on satisfaction with the police, Weitzer and Tuch (2005) showed that public confidence in police was strongly predicted by the respondents' perception of effective crime control. Similar evidence was found by Dekker and Van der Meer (2007). Adding performance indicators to their model of confidence in the Dutch criminal justice system tripled explained variance. *Performance* can for instance be measured by asking respondents how often the police provide satisfactory service, and how well the courts solve problems (Tyler, 2001). Another way of gauging performance, applied in the British Market & Opinion Research International (MORI) 2003 poll, is by asking respondents to express how much confidence they have in specific functions of the criminal justice system. The functions which respondents were asked about included "stopping offenders from committing more crime" and "creating a society in which people feel safe" (Roberts & Hough 2005). Another example is the International Crime Victims Survey (ICVS), which asks respondents how good a job the police do in controlling crime (Van Dijk, Van Kesteren & Smit 2007).

Why measuring procedural justice and performance in the criminal justice system by actor?

Who or what is receiving a particular level of confidence? As mentioned above, the *object* of confidence differs substantively between surveys. In some surveys the objects of interest are the police and the justice system, while in other surveys items refer to the courts, judges, the Supreme Court, or plainly 'justice'. Such object-differentiation can be an important cause for the widely varying confidence levels that emerge from different surveys. As previous studies have clearly demonstrated, a distinction should be made between questioning respondents about the criminal justice system in general, and about specific professionals within that system (Haas, 2010). This distinction is essential because when citizens are asked about the whole system, they may provide an answer with a specific actor in mind (Dekker & Van der Meer, 2007). Specifying confidence for every actor results in differential confidence ratings, which we have demonstrated earlier (Van Damme et al, 2010). Agencies that are responsible for punishment of offenders, such as courts and the prison system, are likely to receive lower levels of confidence. This 'hierarchy of confidence' is found in most western countries in which respondents rate the performance of specific actors (Roberts & Hough, 2005). To explain the difference in rate performances by actors, Hough and Roberts (2004) maintain that public confidence tends to be higher for those actors in the justice system whose function is closest to the view of criminal justice that most members of the public hold, i.e. the crime control model. Nevertheless, it has not been explicitly studied in Belgium from a methodological point of view.

Reliability and Correlational validity as criteria for the assessment of a measurement instrument

Multi-item assessment instruments, such as scales derived from questionnaires, are frequently used by criminologists. A main reason for this is that they provide converging pieces of information about latent traits or attitudes under investigation. By making use of multiple rather than single indicators of unobservable constructs, one is in a far better position to discover relationships between constructs. With such multiple-measure instruments, the question about their reliability asks for the degree to which the final score on the composite (e.g. the sum of their components) is affected by error. To the extent that an instrument's reliability is high, one can be confident that its total score reflect genuine individual differences in the latent dimension of concern. Conversely, with low reliability of the involved scales, the researcher is at serious risk of drawing incorrect conclusions about

relationships between latent variables. In particular, even if one has a correct conceptualisation of the relationships among studied constructs (i.e. correctly specifies a model for them), the use of instruments with insufficient reliability can make it more likely that the researcher will not find support for the model in the data, mostly due to poor measurement. Alternatively, the researcher can end up with incorrect parameter points and interval estimates suggesting misleading substantive interpretations. At least as importantly, high reliability of a measurement is a necessary condition for high validity, the bottom line of measurement. It is therefore of particular importance for inquiries of confidence to be concerned with the instruments *reliability* and *validity*. In the present study, *reliability* is the consistency of a measurement; it is restricted to the study of internal consistency of the measurement instrument. Internal consistency estimates reliability by grouping questions that measure the same concept in a questionnaire. One classic way of computing correlation values among the questions in the instrument is by using Cronbach's alpha. The closer the Cronbach's alpha value is to one, the higher the reliability estimate of the instrument. Nevertheless, Cronbach's alpha is a rather conservative way of estimating the reliability of a measurement instrument and may even generate high values when scales are multi-dimensional. In that case, different theoretical aspects are put together without a solid basis for doing so. As a result, correlations between these variables and their presumed causes may be misleading. *Validity* concerns the truth of measurements. There are three common methods to determine the validity of scores. One method is based on logical evaluation by content experts (content-related or face validity). A second method is based on empirical evidence from statistical analysis (criterion-related validity). A third one is based on a complex series of hypothesis tests and statistical analyses (construct-related validity). The present study uses correlational validity as a means of gaining insights into the construct validity of measures of procedural justice and performance of the criminal justice system.

The latent variables approach and the study of non-random error due to common unknown variance

Although Belgium has an ongoing tradition of ten years in measuring confidence in the criminal justice system by means of the “Justitiebarometer” (Parmentier et al, 2004), a latent variable approach to the study of measurement error in inquiries of confidence in the criminal

justice system is rather new in the Belgian context.[1] The present study therefore aims to fill that gap by studying the internal consistency (factor structure) and correlational validity of our measurement instrument. We prefer a latent variable approach as opposed to a classic Likert sum-scale approach because the latter does not allow for understanding the relative importance of an item and simply adds all items to create the scale construct. Some items may contribute more to the latent construct than other items and thus deserve a larger weight. Furthermore, Likert-scales are based on a key assumption of classical test-theory (Lord & Novick, 1968) and thus do not assume correlated errors between indicators or even latent variables in higher-order factor analyses. The latent variable approach is specifically designed to gain insights into measurement models by studying the factor structure. Theoretically relevant measures for complex constructs such as dimensions of confidence in the criminal justice system can be evaluated using the latent variable approach by examining the fit of theoretically specified measurement models. A series of global model fit indices exist for that purpose and will be discussed later. A theoretically acceptable measurement model should normally be free of correlated error terms between indicators (survey questions) (Raykov, 2001, 2004). Some scholars argue that correlated error terms between indicators may indicate a conceptual and theoretical misspecification. Conversely, poor fit, defined as the difference between the predicted and observed indicator covariance in a measurement model, can nearly always be remedied by adding a correlated error term. In practice, correlated measurement errors are too often employed in a post hoc manner to obtain an acceptable fit of the model to the data (Bagozzi 1983; Fornell 1983). We warn against such practices because the post hoc addition of a correlated measurement error simply means that the observed covariance between a given pair of indicators has not been adequately accounted for by the factors present in the original model (Gerbing & Anderson, 1984). While the use of correlated measurement errors improves a model fit by accounting for this unwanted covariance, it does so at a correspondent loss of meaning and substantive conclusions which can be drawn from the model. Their post hoc use means that indicator covariance is due to at least one unknown common source. As correlated measurement terms are added without theoretical guidance, the correspondence between the posited construct of interest and the empirically defined factor

[1] The 'Justitiebarometer' (English: Justice survey) is an official study that was developed in 2000 and is now conducted in the "Hoge Raad voor Justitie" (In English: High Council for Justice) and whose primary goal is to assess different attitudes such as satisfaction with and confidence towards the criminal justice system and different courts. The results of the Belgian Justice survey are restricted to the presentation of cross-tabulations by item. Methodological inquiries into the reliability and validity of its key constructs have never been the goal of the measurement instrument.

becomes unclear. A preferred substantive representation of this error covariance would be to model it separately from the construct of interest, but that is not always possible. It is important to notice that in some situations the use of correlated measurement errors is meaningful and should even be specified a priori (e.g. panel studies in which the same questions are used over time).

We argue that the study of non-random measurement error may also be applied to the study of confidence in the criminal justice system, in cases when similar indicators of procedural justice and performance are posed with regard to different actors in the criminal justice system. Our measurement instrument uses highly similar questions that measure procedural justice and performance for different actors of the criminal justice system and this may lead to error covariance in two ways: (1) between indicators with almost identical formulations and (2) between questions that refer to the same actor. If one does not take the existence of such question wording effects into account, a measurement instrument may be rejected on false grounds. In other words, wrong conclusions may be drawn from the measurement model because the question wording error was not taken into account.

To the best of our knowledge, to what extent a measurement instrument that is dealing with dimensions of confidence by actors is sensitive to correlated errors, has not been studied in detail. Such insights are important because measurement error can disturb the 'true' correlations of both dimensions of confidence and their correlates and the 'true factor loadings' of each indicator with regard to the construct it measures. The present study therefore addresses this gap in the 'confidence literature' by raising the methodological issue of the latent structure of dimensions of confidence in the criminal justice system and the correlations between those different measures. The latent variables approach has a number of strong evaluation criteria in order to compare different hierarchical models. More specifically, it becomes possible to evaluate the different latent variables by checking the improvement in model fit (Billiet & McClendon, 2000). In short, the latent variables approach is necessary to reach the goals of the present study, i.e. gaining insights into the reliability and validity of the measurement instrument in two ways. First of all, the latent variables approach allows us to identify the latent structure of our measurement instrument that measures two dimensions of confidence in the criminal justice system (performance and procedural justice) with regard to three actors of the criminal justice system and the criminal justice system in general. Using

the latent variable approach will give us insights into the way possible question wording effects have an impact on the factor structure and model fit of the theoretically derived multidimensional model of confidence in the criminal justice system. Secondly, the latent variables approach allows us to gain insights into the correlational validity of the latent variables that are derived from the measurement instrument.

Introducing the questionnaire

The measurement instrument used in the present study contains 34 attitude statements. The respondents had to indicate how closely their feelings match the statement on a 5-point Likert scale (totally disagree to totally agree). Table 1 shows the resulting 34 items, which are ordered by the different actors: (criminal) judges, prosecutors, police and the criminal justice system (CJS). The concept column shows the distinction between procedural justice and performance. The items were largely adopted from the questionnaire used by Haas (2010). Haas (2010) made a selection of items that were identified in the literature as indicators of performance and procedural justice to study the effects of confidence on vigilantism during her PhD study.

Table 1: Questionnaire items measuring confidence in the Criminal Justice System

Items	Concept	Actor
V8_1 Judges treat people fairly	**Procedural Justice**	*Judges*
V8_2 The Public Prosecution is honest and trustworthy	**Procedural Justice**	*Public Prosecution*
V8_3 Citizens can count on it that their case is properly dealt with in the Belgian criminal justice system	**Procedural Justice**	*Criminal Justice System*
V8_4 The police are there when you need them	**Procedural Justice**	*Police*
V8_5 Judges are prejudiced	**Procedural Justice**	*Judges*
V8_6 The Public Prosecution is trustworthy	**Procedural Justice**	*Public Prosecution*

V8_7 The Police are trustworthy	**Procedural Justice**	*Police*
V8_8 The Belgian Criminal Justice System is honest	**Procedural Justice**	*Criminal Justice System*
V8_9 Judges are trustworthy	**Procedural Justice**	*Judges*
V8_10 The Public Prosecution treats people fairly	**Procedural Justice**	*Public Prosecution*
V8_11 The Criminal Justice System is trustworthy	**Procedural Justice**	*Criminal Justice System*
V8_12 Citizens can count on it that their case is properly dealt with in the Belgian Criminal Justice System	**Procedural Justice**	*Criminal Justice System*
V8_13 The Police take citizens seriously	**Procedural Justice**	*Police*
V8_14 The police care about the well-being of the everyday citizen	**Procedural Justice**	*Police*
V8_15 You can count on the judges to take decisions that are best for society	**Procedural Justice**	*Judges*
V8_16 You can count on the Public Prosecution to take decisions that are best for society	**Procedural Justice**	*Public Prosecution*
V8_17 You can count on the Police to take decisions that are best for society	**Procedural Justice**	*Police*
V10_2 If the Police do not arrest someone, they will have a valid reason for that	**Procedural Justice**	*Police*
V10_5 If the Public prosecution recommends a lenient sentence, it will have a valid reason for that	**Procedural Justice**	*Public Prosecution*
V10_6 If a Judge passes a lenient sentence, he will have a valid reason for that	**Procedural Justice**	*Judges*
V9_1 Judges' verdicts are well deliberated	**Performance**	*Judges*
V9_2 The Public Prosecution succeeds in prosecuting the right people	**Performance**	*Public Prosecution*
V9_3 Judges know what's going on in society	**Performance**	*Judges*

V9_4 The Belgian criminal justice system functions properly	**Performance**	*Criminal Justice System*
V9_5 Judges do their job well	**Performance**	*Judges*
V9_6 Sentence recommendations are well-deliberated by the Public Prosecution	**Performance**	*Public Prosecution*
V9_7 Citizens' rights are not protected well by the Public Prosecution	**Performance**	*Public Prosecution*
V9_8 The Belgian justice system is effective in combating crime	**Performance**	*Criminal Justice System*
V9_9 the Public Prosecution does its job well	**Performance**	*Public Prosecution*
V9_10 The Police are effective in combating crime	**Performance**	*Police*
V9_11 The Belgian Criminal Justice system succeeds in bringing criminals to justice	**Performance**	*Criminal Justice System*
V9_12 The Police do a good job	**Performance**	*Police*
V9_13 Citizens' rights are not protected well by judges	**Performance**	*Judges*
V9_14 In the Belgian Criminal Justice System, too much emphasis is put on offenders' rights	**Performance**	*Criminal Justice System*

Data

After describing the statistical framework for our analysis of the reliability and validity of the measurement instrument, and the study of random and non-random measurement error, we now return to the survey data that were used for the analyses. During the academic year 2009-2010, a questionnaire was administered to 1,079 university students. This survey assesses perceived procedural justice of the criminal justice system and its correlates, such as anomia, authoritarianism, ethnocentrism, punitiveness and support for vigilantism. The data collection was carried out within the framework of a compulsory course in quantitative methods taught in the second year of the bachelor degree in the criminological sciences at Ghent University. Students were asked to recruit other students to fill in the questionnaire. The academic staff provided the students with clear instructions on how the data had to be collected (self-administered questionnaires). The students were told that the questionnaire was part of a pre-test of a larger survey (a PhD study in criminology). The main aim of the study was to test

the reliability and correlational validity of the scales used. It was especially instructed that every student should conduct seven interviews and it was stressed that every student should contact a wide variety of students, i.e. students that were enrolled at different faculties. It was stressed that homogeneous samples were to be avoided. Students were therefore only allowed to interview one other student enrolled in the criminological sciences. We guaranteed full confidentiality. The interviewees could fill in the questionnaire and deliver it to the students after completion. We are aware that there may be a selection bias at the interviewer level, but we were able to monitor this through the entire course. Afterwards a multilevel analysis was performed to estimate the interviewer variance with regard to the scales that were constructed. The results suggest that the interviewer variance was negligible, based on the study of the intra-class coefficients per interviewer.[2] 22.3% of the respondents were students of criminology at the faculty of law. 13.3% of the respondents were studying at the faculty of psychology and 10.6% of the respondents were enrolled at the faculty of Arts and Humanities (including philosophy, language and literature sciences). The 53.8 % other respondents studied law, science, social and political sciences, medical science, bio-engineering, pharmaceutical sciences, engineering, economic sciences or veterinary.[3] 43.6% of the respondents were male, 56.4% of the respondents were female. On average, students were 20.2 years old (std= 1.77).

Analysis plan

Confirmatory factor analysis (CFA) is a statistical technique used to verify the factor structure of a set of observed variables. CFA allows the researcher to test the hypothesis that a relationship between observed variables and their underlying latent constructs exists. The researcher uses knowledge of the theory, empirical research, or both, postulates the relationship pattern a priori and then tests the hypothesis statistically. In order to test the latent structure of the confidence-items structural equation modelling was used. All analyses were carried out using LISREL 8.53 (Jöreskog & Sörbom 2003). All models presented in this contribution show standardised factor solutions, and for the evaluation and fit of the models the Root Mean Square Error of Approximation (RMSEA) is preferred over the CHI² value because large samples inevitably tend to result in large CHI² values, and therefore in a

[2] The intra-class coefficients were below 1% with an average of seven interviews per interviewer and thus standard errors obtained from single-level analyses cannot be biased.
[3] For a more detailed overview, see appendix

negative evaluation of the fit of the model. As the RMSEA is least affected by the sample size, it is a good measure of close fit (Mueller, 1996). RMSEA values below .05 point at a good fit of the model (Billiet & McClendon, 2000).

Results

In this paragraph we present the results of a series of confirmatory factor analyses. To assess the aforementioned issue of non-random measurement error at the item and actor level, a series of research questions serve as guidelines:

Research question 1a: Do the data support a non-random error eight-factor model that distinguishes between procedural justice of judges, procedural justice of the police, procedural justice of the public prosecution, procedural justice of the criminal justice system in general, performance of judges, performance of the police, performance of the public persecution and performance of the criminal justice system in general? *Research question 1b*: Do the data support the idea that the actor-specific dimensions of procedural justice and performance actually reflect procedural justice and performance as higher-order latent variables?

Research question 2: Do the data support a random error eight-factor model that distinguishes between procedural justice of judges, procedural justice of the police, procedural justice of the public prosecution, procedural justice of the criminal justice system in general, performance of judges, performance of the police, performance of the public persecution and performance of the criminal justice system in general when non-random error between latent variables that refer to the same actor are taken into account?

Our two research questions can be visualised as follows:

Figure 1: the random error factor model of confidence in the Criminal Justice System

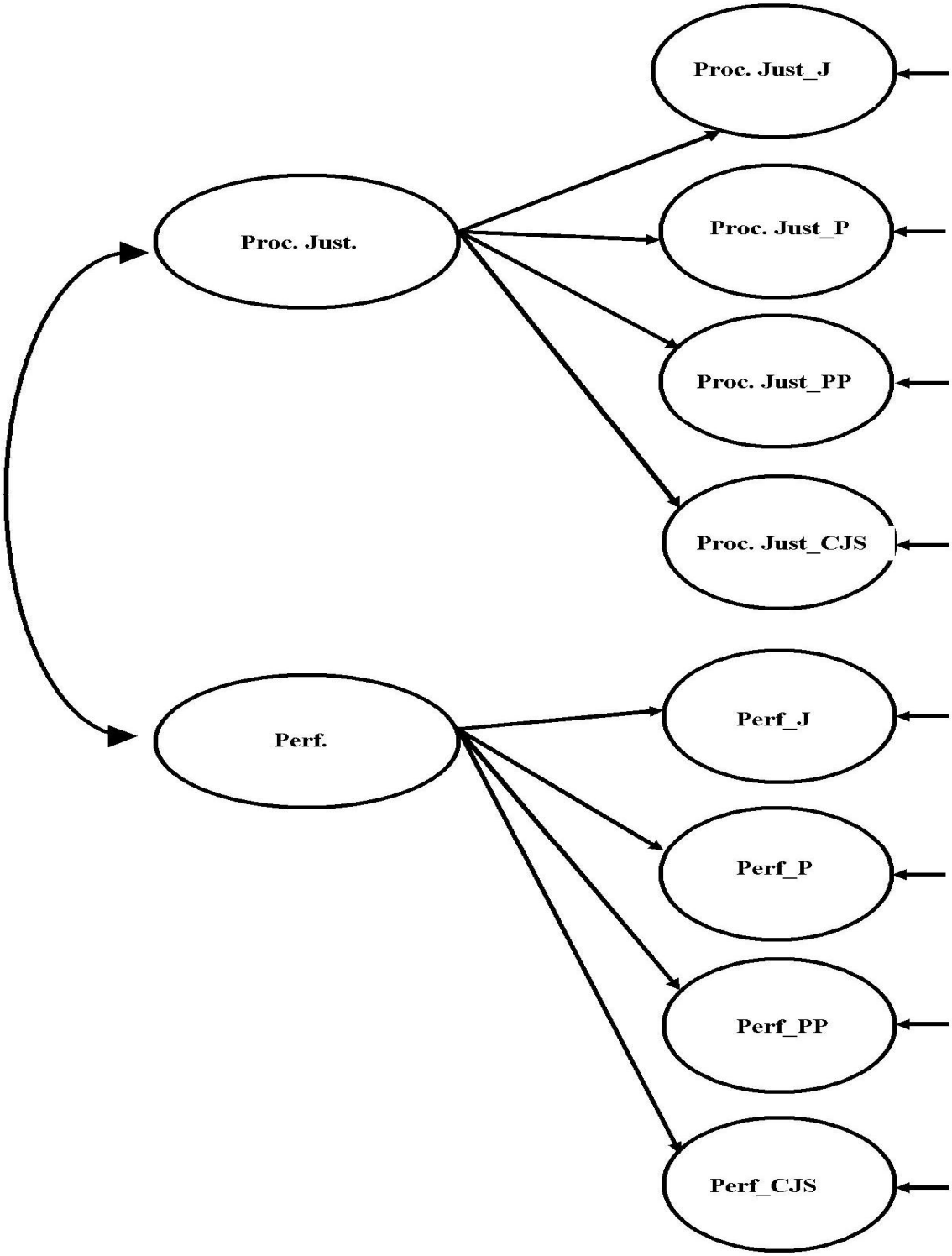

Figure 2: the non-random error factor model of confidence in the Criminal Justice System

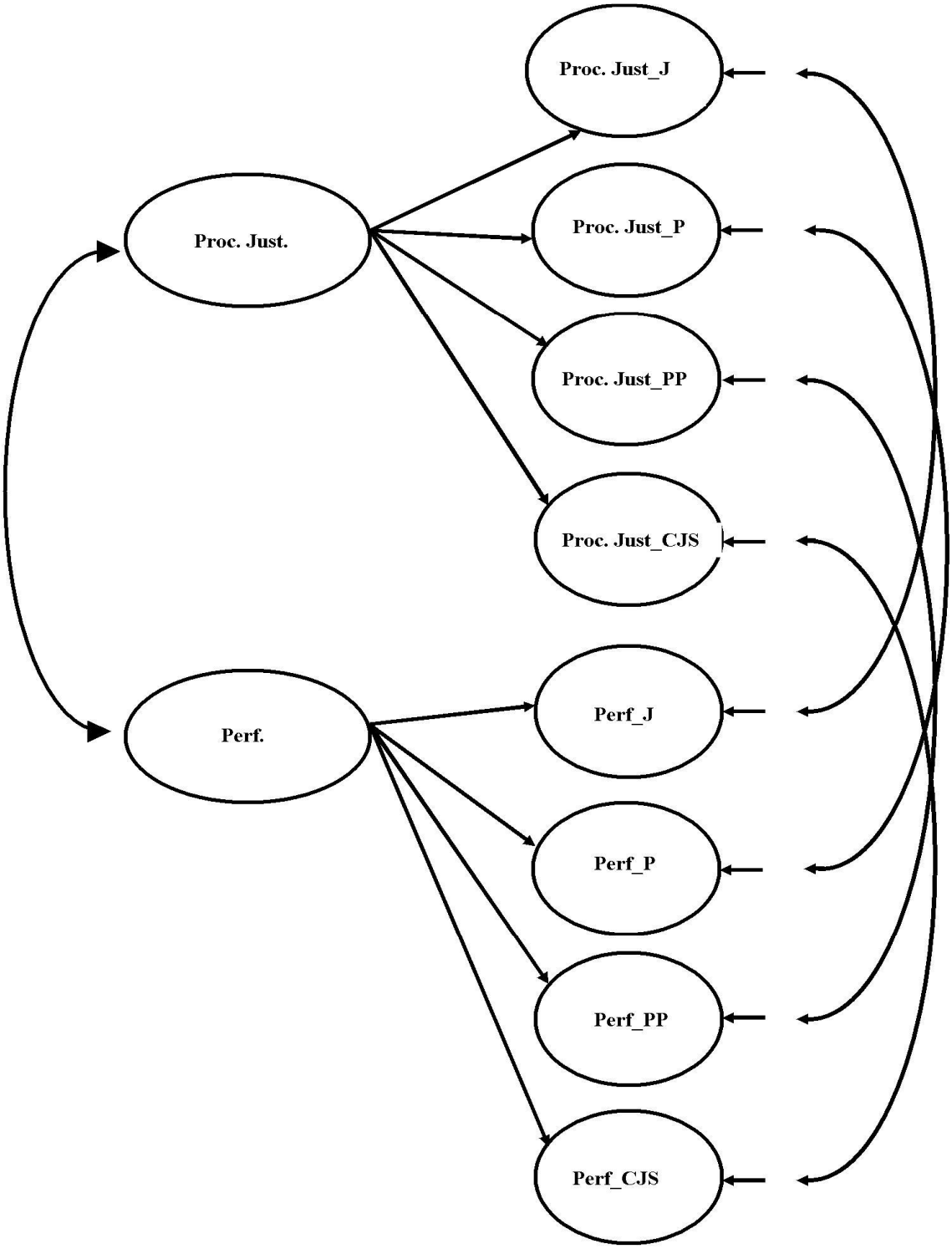

Table 2a: Factor structure of 8-factor model of confidence by dimensions and actor (non-random error model)-results at the item level

Items	Proc Just_J	Proc Just_P	Proc Just_PP	Proc Just_CJ	Perf_J	Perf_P	Perf_PP	Perf_CJ
V8_1	0.75	--	--		--	--	--	--
V8_2	--	--	0.56		--	--	--	--
V8_3	--	--	--	0.79	--	--	--	--
V8_4	--	0.74	--		--	--	--	--
V8_5 (r)	0.57	--	--		--	--	--	--
V8_6 (r)	--	--	0.85		--	--	--	--
V8_7	--	0.83	--		--	--	--	--
V8_8	--	--	--	0.80	--	--	--	--
V8_9	0.82	--	--		--	--	--	--
V8_10 (r)	--	--	0.85		--	--	--	--
V8_11	--	--	--	0.86	--	--	--	--
V8_12	--	--	--	0.83	--	--	--	--
V8_13	--	0.81	--	--	--	--	--	--
V8_14	--	0.76	--	--	--	--	--	--
V8_15	0.61	--	--	--	--	--	--	--
V8_16 (r)	--	--	0.64	--	--	--	--	--
V8_17	--	0.73	--	--	--	--	--	--
V9_1	--	--	--	--	0.74	--	--	--
V9_2	--	--	--	--		--	0.65	--
V9_3	--	--	--	--	0.47	--	--	--
V9_4	--	--	--	--		--	--	0.82
V9_5	--	--	--	--	0.83	--	--	--
V9_6	--	--	--	--	--	--	0.64	--
V9_7(r)	--	--	--	--	--	--	0.43	--
V9_8	--	--	--	--	--	--		0.77
V9_9	--	--	--	--	--	--	0.83	--
V9_10	--	--	--	--	--	0.83	--	--
V9_11	--	--	--	--	--		--	0.78
V9_12	--	--	--	--	--	0.89	--	--
V9_13 (r)	--	--	--	--	0.62	--	--	--

V9_14(r)	--	--	--	--	--	--	--	0.31
V10_2	--	0.37	--	--	--	--	--	--
V10_5 (r)	--		0.47	--	--	--	--	--
V10_6	0.55			--	--	--	--	--

Chi-Square=6387.89, df=518, P-value=0.00000, RMSEA=0.103 AGFI = 0.70

From the values in Table 2a it can be seen that respondents are able to make a distinction between actors and dimensions of confidence. We are able to make a distinction between procedural justice of the criminal justice system in general, performance of the criminal justice system in general, procedural justice of the public prosecutor, performance of the public prosecutor, procedural justice of judges, performance of judges, procedural justice of the police and performance of the police. All factor loadings are statistically significant from zero.

Table 2b: Factor structure of the second order factors (non-random error model)

First-order factors	Procedural justice	Performance
ProcJust_J	0.91	
ProcJust_P	0.67	
ProcJust_PP	0.87	
ProcJust_CJ	0.86	
Perf_J		0.90
Perf_P		0.61
Perf_PP		0.92
Perf_CJ		0.77

The higher-order factor analysis suggests that all actor-specific confidence measures have substantial and significant loadings on a general confidence second-order latent variable. It is worthwhile to notice that confidence in the police has the lowest factor loading.

Table 2c: correlation matrix between latent variables (random error model)

	Proc Just_J	Proc Just_P	Proc Just_PP	Proc Just_CJS	Perf _J	Perf _P	Perf _PP	Perf _CJS
ProcJust_J	1							
ProcJust_P	0.61	1						
ProcJust_PP	0.79	0.59	1					
ProcJust_CJS	0.78	0.58	0.75	1				
Perf_J	0.80	0.59	0.77	0.75	1			
Perf_P	0.55	0.40	0.53	0.52	0.55	1		
Perf_PP	0.82	0.61	0.79	0.78	0.83	0.57	1	
Perf_CJS	0.69	0.51	0.66	0.65	0.69	0.47	0.71	1

All correlations significant at $p < 0.01$ or better

From the correlation matrix between the latent constructs it can be seen that actor-specific procedural justice and actor-specific performance are highly correlated. On the other hand, it is found that respondents that positively evaluate the procedural justice or performance of one actor, also are more likely to do so for the other actors. All correlations are substantive and statistically significant. It is striking that the latent constructs that refer to the criminal justice system, judges and public prosecution are correlated more highly with each other than with the latent constructs that refer to the police. All correlations are in the expected direction. Interestingly and in line with previous studies, it was found that those that have high levels of confidence in one actor, also seem to have high levels of confidence in another actor. The correlations between the levels of procedural justice and performance of the police and the same constructs with regard to the other actors of the criminal justice system (prosecution, judges, criminal justice system in general) are somewhat lower than the correlations between the dimensions of confidence within all dimensions of the criminal justice system. This is possibly due to the fact that the police is considered different from the other actors of the criminal justice system. Our findings suggest that it is of the utmost importance to take these differences into account when the confidence is measured. The two higher order factors are correlated 0.98 in this model which does not take non-random errors into account. This is surprising. It suggests that there is virtually no empirical distinction between procedural justice and performance. It is difficult to explain. There is no reason to assume that a student

population is less able to make a distinction between both dimensions than respondents in a general population. In fact, we would have expected that students are more able to distinguish both constructs than respondents in a general population. The general model fit of this higher-order factor model is far from perfect (RMSEA: 0.10 and AGFI: 0.70). The main reason, as suggested by the modification indices is that to obtain an adequate model fit, all too many error covariances need to be freed.[4] From a close inspection of these modification indices, we learn that virtually all items that refer to a same actor should be freed to obtain an acceptable model fit. This approach is not always accepted by scholars, as it is often done without serious theoretical motivation, which leads to freeing error covariances that cannot be justified. In that case any result obtained is merely data-driven. Alternatively, we hypothesised that there may be two distinct sources of error covariance: (1) error covariance at the latent variable level, due to the fact that we measure procedural justice and performance towards four actors and (2) error covariance at the item level resulting from high similarities in question wording. This means that we may expect four error covariances at the latent variable level: performance of the police may share an error covariance with procedural justice of the police, performance of judges may share an error covariance with procedural justice of judges, performance of the criminal justice system may share an error covariance with procedural justice of the criminal justice system and performance of the public prosecutor may share an error covariance with procedural justice of the public prosecutor.[5]

Table 3a: Factor structure of the 8-factor model of confidence by dimensions and actor (random error model at the indicator level and first order factor level)

Items	Proc Just_J	Proc Just_P	Proc Just_PP	Proc Just_CJS	Perf _J	Perf _P	Perf _PP	Perf _CJS
V8_1	0.75	--	--		--	--	--	--
V8_2	--	--	0.60		--	--	--	--
V8_3	--	--	--	0.81	--	--	--	--

[4] The modification index is a lower bound estimate of the expected chi square decrease that would result when a particular parameter is left unconstrained (making it a free parameter, or adding it as an extra path). Jöreskog suggested that a modification index should be at least five before the researcher considers modifying the hypothesised model.

[5] In the analyses therefore error covariances were released between the latent variables ProcJust_J and Perf_J, ProcJust_P and perf_P, ProcJust_PP and Perf_PP, ProcJust_CJS and Perf_CJS and between following indicators: V8_5 and V8_2, V8_7 and V8_6, V8_14 and V8_13, V8_16 and V8_15, V8_17 and V8_16, V9_10 and V9_8, V10_5 and V10_2, V10_6 and V10_2, V10_6 and V10_5 (see Table 1 for the content of this items).

V8_4	--	0.74	--		--	--	--	--
V8_5 (r)	0.59	--	--		--	--	--	--
V8_6 (r)	--	--	0.85		--	--	--	--
V8_7	--	0.82	--		--	--	--	--
V8_8	--	--	--	0.81	--	--	--	--
V8_9	0.82	--	--		--	--	--	--
V8_10 (r)	--	--	0.87		--	--	--	--
V8_11	--	--	--	0.87	--	--	--	--
V8_12	--	--	--	0.85	--	--	--	--
V8_13	--	0.77	--	--	--	--	--	--
V8_14	--	0.70	--	--	--	--	--	--
V8_15	0.57	--	--	--	--	--	--	--
V8_16 (r)	--	--	0.62	--	--	--	--	--
V8_17	--	0.70	--	--	--	--	--	--
V9_1	--	--	--	--	0.73	--	--	--
V9_2	--	--	--	--		--	0.65	--
V9_3	--	--	--	--	0.47	--	--	--
V9_4	--	--	--	--		--	--	0.82
V9_5	--	--	--	--	0.83	--	--	--
V9_6	--	--	--	--	--	--	0.64	--
V9_7(r)	--	--	--	--	--	--	0.43	--
V9_8	--	--	--	--	--	--		0.77
V9_9	--	--	--	--	--	--	0.83	--
V9_10	--	--	--	--	--	0.78	--	--
V9_11	--	--	--	--	--		--	0.78
V9_12	--	--	--	--	--	0.88	--	--
V9_13(r)	--	--	--	--	0.61	--	--	--
V9_14(r)	--	--	--	--	--	--	--	0.31
V10_2	--	0.37	--	--	--	--	--	--
V10_5(r)	--		0.38	--	--	--	--	--
V10_6	0.43			--	--	--	--	--

Chi-Square=3146.31, df=507, P-value=0.00000, RMSEA=0.06 AGFI: 0.84

All hypothesised non-random error covariances are substantial (standardised solution: 0.2-0.3). The non-random measurement model does not seem to affect the factor loadings of the latent variable model, while the model fit has improved substantively.[6] For that reason, we do not discuss this factorial structure more precisely.

Table 3b factor structure of the second order factors (non-random error model)

First-order factors	Procedural justice	Performance
ProcJust_J	0.88	
ProcJust_P	0.65	
ProcJust_PP	0.86	
ProcJust_CJ	0.88	
Perf_J		0.87
Perf_P		0.57
Perf_PP		0.57
Perf_CJ		0.94

Interestingly, taking expected error covariances at the latent variable level (actors) into account and item level simultaneously, reveals some substantive changes in the loadings of the latent variables towards the second-order latent variables. Procedural justice of the police remains the latent variable with the lowest loading on procedural justice. With regard to the performance of the criminal justice system, the performance of the police and the public prosecutor have now both the lowest factor loadings.

[6] It is important to notice that the model fit can be improved beyond our methodological exercise. However, in that case we would merely follow the structure of the data and that would be against the philosophy behind SEM modelling.

Table 3b: correlation matrix between latent variables (non-random error model)

	Proc Just_J	Proc Just_P	Proc Just_PP	Proc Just_CJS	Perf _J	Perf _P	Perf _PP	Perf _CJS
ProcJust_J	1							
ProcJust_P	0.58	1						
ProcJust_PP	0.76	0.56	1					
ProcJust_CJS	0.78	0.58	0.76	1				
Perf_J	0.92	0.50	0.67	0.68	1			
Perf_P	0.45	0.90	0.44	0.44	0.49	1		
Perf_PP	0.74	0.55	0.93	0.74	0.82	0.53	1	
Perf_CJS	0.63	0.65	0.62	0.82	0.70	0.45	0.75	1

All correlations significant at $p < 0.01$ or better

The correlations of these different dimensions of procedural justice and performance have changed marginally, but in general rather similar to the ones presented in the Table 2c. The higher order factors (procedural justice and performance) have a correlation of 0.89, when taking non-random error into account.

Discussion and conclusion

The most important question of this exercise remains the following: How do we know which set of estimates to believe? Given the circumstances in which the questions were asked as follows: common question format, common response format (5-point scales), each question in immediate proximity to the others -there are strong theoretical reasons for suspecting non-random error to come into play. The difference in chi-square of chi-square model 1 (6871.89) minus chi-square model 2 (3146.31) and the difference in degrees of freedom (518-507=11) is significant at $p < 0.001$. The non-random error model fits the data better. This statistical test leaves little doubt that the non-random error model provides a superior fit to the data. We now turn to the answers of the research questions posed:

Research question 1a: Do the data support a non-random error eight-factor model that distinguishes between procedural justice of judges, procedural justice of the police, procedural justice of the public prosecution, procedural justice of the criminal justice system

in general, performance of judges, performance of the police, performance of the public persecution and performance of the criminal justice system in general? Research question 1b: Do the data support the idea that the actor-specific dimensions of procedural justice and performance actually reflect procedural justice and performance as higher-order latent variables?

The data do not support the idea that a non-random error model of procedural justice and performance leads to an acceptable model fit, although all factor loadings were statistically significant. This means that procedural justice and performance can be measured with regard to different actors. In the non-random model the correlation between all latent variables were significant, and the correlation between both higher-order factors was very high (0.91).

Research question 2: Do the data support a random error eight-factor model that distinguishes between procedural justice of judges, procedural justice of the police, procedural justice of the public prosecution, procedural justice of the criminal justice system in general, performance of judges, performance of the police, performance of the public persecution and performance of the criminal justice system in general when account is taken for non-random error at the actor level and at the item level?

The data suggest that the non-random error model fits the data much better than the random error model. We identified error covariances amongst the items due to item similarities (question wording). Due to the fact that we measured two dimensions of confidence (procedural justice and performance) for each and every actor, error covariances were specified at the latent variable level. Taking non-random error into account that may arise from question wording and from the fact that measuring different dimensions of confidence by actor, improves the model fit. One may argue that the model fit improves by definition because a non-random error model poses less restriction to the data than a random error model. However, the non-random error that has been identified, was specified and we provided a theoretical rational for the non-random error. The good news is that there was almost no difference in the correlation between the latent variables and the factor loadings remained similar. In other words, correlational validity and reliability of the measurement are not affected in a problematic way.

What lessons can we draw from this methodological exercise? A common criticism of methodological exercises is that there is little surplus value to gain when trying to model the complex relations between indicators of confidence in the criminal justice system (Pauwels & Pleysier, 2005). Yet, without studying the factor structure, it is impossible to properly understand why some models have an appropriate model fit and some other models don't have an appropriate model fit. Methodological exercises to gain insights into the factorial structure are important for studies that want to address issues concerning the covariates of confidence from competing theoretical perspectives. Before studying the theoretical correlates of procedural justice and performance, the quality of these latent variables should be tested. This study reaffirms that it is possible to measure different dimensions of confidence by actors in one questionnaire. The amount of error covariance between error covariance is huge when using a detailed questionnaire. Indeed, when providing interviewees with highly similar items for every construct, non-random measurement error is due to be a reality. If one does not model that reality, the true correlations between all latent constructs may be affected by not taking these sources of measurement error into account. However, from the present study it became clear that, while non-random error exists, the correlations between the different latent variables were rather unaffected.

What are the implications of the current findings for future studies of confidence in the criminal justice system? How detailed the focus should be with regard to measurement issues depends on the objective of the study. If the objective of a survey is to evaluate the respondents' attitudes towards specific aspects of the criminal justice system, then the items that measure both dimensions of confidence by actor are recommended. Our study reaffirms that this is possible.

What does our analysis mean with regard to previous studies? We are not saying that any of the vast number of studies that have applied conventional techniques of assessment to commonly worded batteries of survey items have wrongly estimated the quality of their measures. Our aim was rather to alert scholars to the symptoms of non-random error and to suggest that researchers test their substantive claims after a thorough analysis of the latent structure of the measures to be used. To do so, it is necessary to make allowances for non-random error when constructing survey instruments, so that statistical models of inter-item relationships can incorporate adjustments for non-random error.

Appendix

Background characteristics of the respondents

Table: % of the respondents by study

%	Studies
22.3	Criminological sciences
13.3	Psychology and educational sciences
10.6	Humanities and arts
9.5	Medical science
9.2	Social and political sciences
8.6	Law
7.9	Economic sciences
6.7	Engineering
3.7	Science
2.7	Bio-engineering
2.1	Pharmaceutical sciences
1.8	other
1.5	Veterinary sciences
100 %	**Total respondents (N=1101)**

Table: Faculty of the respondents

Faculty

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Law faculty	340	30,9	30,9	30,9
	social sciences	364	33,1	33,1	63,9
	science & meds	290	26,3	26,3	90,3
	economics	87	7,9	7,9	98,2
	Univ college	20	1,8	1,8	100,0
	Total	1101	100,0	100,0	

Table: Gender of the respondents

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	480	43,6	43,6	43,6
	Female	621	56,4	56,4	100,0
	Total	1101	100,0	100,0	

References

Bagozzi, R. (1983). Issues in the Application of Covariance Structure Analysis: A Further Comment, *Journal of Consumer Research*, 9 : 449-450.

Benesh, S. C. & Howell, S. (2001). Confidence in the Courts: A Comparison of Users and Non-Users. *Behavioral Sciences and the Law,* 19, 199-214.

Benesh, S. C. (2006). Understanding public confidence in American courts. *The Journal of Politics, 68*(3), 679-707.

Billiet, J.B. & McClendon, J. (2000). Modeling Acquiescence in Measurement Models for Two Balanced Sets of Items. Structural Equation Modeling,. 7, 4, 608-628.

Caldeira, G. A. & Gibson, J. L. (1995). The legitimacy of the court of justice in the European Union: models of institutional support. *American Political Science Review, 89*(2), 356-376.

Dekker, P. & Van der Meer, T. (2007). *Vertrouwen in de rechtspraak nader onderzocht*. Den Haag: Sociaal en Cultureel Planbureau.

Fornell, C. (1983). Issues in the Application of Covariance Structure Analysis: A Comment, *Journal of Consumer Research*, 9: 443-448.

Gau, J. (2011). The Convergent and Discriminant Validity of Procedural Justice and Police Legitimacy: An Empirical Test of Core Theoretical Propositions. *Journal of Criminal Justice* 39: 489–498.

Genn, H. & Paterson, A. (2001). *Paths to justice Scotland. What people do and think about going to law*. London: Hart Publishing.

Gerbing, D.W. & Anderson, J.C. (1984). On the meaning of within factor correlated measurement errors. *Journal of consumer research*, 11, 572-580.

Haas, N. (2010). *Public Support for Vigilantism*. Doctoraal Proefschrift. Universiteit Leiden.

Hough, M. & Roberts, J.V. (2004) *Confidence in Justice: an International Review*. London: King's College.

Hough, M., Jackson, J., Bradford, B., Myhill, A. and Quinton, P. (2010). 'Procedural Justice, Trust and Institutional Legitimacy', *Policing: A Journal of Policy and Practice*, 4, 3, 203-210.

Jackson, J. & Sunshine, J. (2007). Public Confidence in Policing: A Neo-Durkheimian Perspective, *British Journal of Criminology,* 47, 2, 214-233.

Jackson, J., Bradford, B., Hough, M., Kuha, J., Stares, S. R., Widdop, S., Fitzgerald, R., Yordanova, M. and Galev, T. (2011). 'Developing European Indicators of Trust in Justice', *European Journal of Criminology*, 8, 4, 267-285.

Jöreskog & Sörbom (2003) Lisrel 8: Structural equation Modeling with the SIMPLIS Command Language, Chicago: SSI.

Lord, F.M, & Novick, M.R. (1968). *Statistical theories of mental test scores.* Reading Mass.: Addison-Wesley.

Mueller, R.A. (1996). Basic principles of structural equation modeling, an introduction to LISREL and EQS. New York: Springer-Verlag.

Pauwels, L. & Pleysier, S. (2005). Effecten Van Antwoordstijlen In Etiologisch Self-Report-Onderzoek, Een 'Causale Modellenbenadering', *Tijdschrift Voor Criminologie*, 47, 1: 42-61.

Pauwels, L. & Pleysier, S. (2009). *Criminaliteit en onveiligheid meten: de gestandaardiseerde vragenlijst.* Leuven: Acco.

Parmentier, S., Vervaeke, G., Goethals, J., Doutrelepont, R., Kellens, G., Lemaître, A., Cloet, B., Schoffelen, J., Vanderhallen, M., Biren,P., Sintobin, M., Van Win, T. & Vandekeere, M., (2004). *Justitie doorgelicht. De resultaten van de eerste Belgische 'justitiebarometer',* Gent, Academia Press.

Raykov, T. (2001). Bias of Cronbach's coefficient alpha for fixed congeneric measures with correlated errors. *Applied Psychological Measurement*, 25, 69-76.

Raykov, T. (2004). Behavioral Scale Reliability and Measurement Invariance Evaluation Using Latent Variable Modeling, *Behavior Therapy,* 35,299—331

Reisig, M.D., Bratton, J. & Gertz, M.G. (2007) 'The Construct Validity and Refinement of Process-based Policing Measures', *Criminal Justice and Behaviour* 34: 1005–27.

Roberts, J. V. & Hough, M. (2005). Public confidence in the criminal justice system, *Understanding public attitudes to criminal justice* (pp. 29-51). Maidenhead: Open University Press.

Sunshine, J., & Tyler, T.R. (2003a). Moral solidarity, identification with the community, and the importance of procedural justice. *Social Psychology Quarterly*, 66, 153-165.

Sunshine, J. & Tyler, T. R. (2003b). The role of procedural justice and legitimacy in shaping public support for policing. *Law & Society Review, 37*(3), 513-548.

Thibaut, J. & Walker, L. (1975). *Procedural justice: A psychological analysis*. Hillsdale: Erlbaum.

Tyler, T. R. (2001). Public trust and confidence in legal authorities: What do majority and minority group members want from the law and legal institutions? *Behavioral Sciences and the Law, 19*, 215-235.

Tyler, T.R. (2003) 'Procedural justice, legitimacy and the effective rule of law', *Crime & Justice*, 30, 283 – 357.

Tyler, T.R. (2006). *Why people obey the law.* New Jersey: Princeton University Press.

Tyler, T.R. (ed.) (2007). *Legitimacy and criminal justice: international perspectives*. New York: Russell Sage Foundation Press.

Tyler, T.R. (2011). *Why people cooperate: The role of social motivations.* Princeton: Princeton University Press.

Van Damme, A., Pauwels, L., Pleysier, S. & Van De Velde, M. (2010). Beelden van vertrouwen: het vertrouwen in politie en justitie in perspectief geplaatst (locomotieftekst). *Orde van de dag* – themanummer: "Beelden van vertrouwen in politie en justitie", 52, 7 – 20.

Van de Velde, M. &Pauwels, L. (2010). Institutional distrust in Flanders. What is the role of social capital and dimensions of discontent? In: Cools, M, De Ruyver, B., Easton, M., Pauwels, L., Ponsaers, P., Vande Walle, G., Vander Beken, T., Vander Laenen, F., Vermeulen, G., Vynckier, G. (Eds), *Governance of Security Research Papers Series III, New Empirical Data, Theories and Analyses on Security, Societal Problems and Citizens' Perceptions*, Antwerpen: Maklu.

Van Dijk, J. J. M., Van Kesteren, J. N. & Smit, P. (2007). *Criminal victimisation in international perspective. Key findings from the 2004-2005 ICVS and EU ICS.* The Hague: Boom legal publishers.

Weitzer, R. & Tuch, S. A. (2005). Determinants of public satisfaction with the police. *Police Quarterly, 8*(3), 279-297.